

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Celeste, Anthony J.
Wozney, John
Rosen, Vicki A.
Wolfman, Neil
Thomsen, Gerald H.
Melton, Douglas A.

(ii) TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS

(iii) NUMBER OF SEQUENCES: 35

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: GENETICS INSTITUTE, INC.
(B) STREET: 87 Cambridge Park Drive
(C) CITY: Cambridge
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Patent In Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Lazar, Steven R.
(B) REGISTRATION NUMBER: 32,618
(C) REFERENCE/DOCKET NUMBER: 5202-D

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617 498-8260
(B) TELEFAX: 617 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 926 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:

(B) CLONE: v1-1

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 571..882

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..882

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCG CGT AAT ACG ACT CAC TAT AGG GCG AAT TGG GTA CGG GGC CCA GGC Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly -190 -185 -180 -175	48
AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC CCG GGC GCC Ser Trp Thr Ser Pro Pro Leu Leu Leu Ser Thr Cys Pro Gly Ala -170 -165 -160	96
GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG CCC CTA GTC Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val -155 -150 -145	144
GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG AGG CGC CAC Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His -140 -135 -130	192
CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG CTG CGC GCA Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala -125 -120 -115	240
GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA CTG GGC TTC Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe -110 -105 -100 -95	288
GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC GCG GTG CTA Gly Trp Pro Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu -90 -85 -80	336
GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC CGG GAG ATC Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile -75 -70 -65	384
CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA GAG CCG CTG Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu -60 -55 -50	432
CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC ATT GGC GGC Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly -45 -40 -35	480
CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA GCG CAG GGC Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly -30 -25 -20 -15	528
AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC CGG AGC CGC Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg -10 -5 1	576
TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC GGC TGG GAC Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp 5 10 15	624
GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC TGC GAG GGC Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly 20 25 30	672
CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC AAC CAT GCC Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala 35 40 45 50	720

ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG GCG CCG GCC	768
Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala	
55 60 65	
TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC CTC TAC ATC	816
Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile	
70 75 80	
GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC ATG GTG GTG	864
Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val	
85 90 95	
GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG GGGCAGCCAC	912
Glu Ala Cys Gly Cys Arg	
100	
GCGGCCGAGG ATCC	926

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 294 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly	-190	-185	-180	-175
Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala	-170	-165	-160	
Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val	-155	-150	-145	
Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His	-140	-135	-130	
Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala	-125	-120	-115	
Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe	-110	-105	-100	-95
Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu	-90	-85	-80	
Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile	-75	-70	-65	
Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu	-60	-55	-50	
Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly	-45	-40	-35	
Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly	-30	-25	-20	-15
Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg	-10	-5	1	

Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
 5 10 15
 Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
 20 25 30
 Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
 35 40 45 50
 Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
 55 60 65
 Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
 70 75 80
 Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
 85 90 95
 Glu Ala Cys Gly Cys Arg
 100

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1207 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:

- (B) CLONE: MP52

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 845..1204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCCG	50
GACTGTGACC CCAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT	100
CCCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA	150
GCCGTTTCGC CCACCCCCCA TCACACCCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT	200
GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA	250
CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA	300
GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTCCG	350
GATCTTGCGG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGGAG GCGGGCGGGC	400
TGCCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CCGGCCTCCT TGCTGGATGT	450
GCGCTCCGTG CCAGGCCTGG ACGGATCTGG CTGGGAGGTG TTCGACATCT GGAAGCTCTT	500
CCGAAACTTT AAGAACTCGG CCCAGCTGTG CCTGGAGCTG GAGGCCTGGG AACGGGGCAG	550
GGCCGTGGAC CTCCGTGGCC TGGGCTTCGA CCGCGCCGCC CGGCAGGTCC ACGAGAGGC	600

CCTGTTCTCTG GTGTTTGGCC GCACCAAGAA ACGGGACCTG TTCTTTAATG AGATTAAGGC	780
CCGCTCTGGC CAGGACGATA AGACCGTGTA TGAGTACCTG TTCAGCCAGC GGCGAAAACG	840
GCGG GCC CCA CTG GCC ACT CGC CAG GGC AAG CGA CCC AGC AAG AAC CTT	889
Ala Pro Leu Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu	
1 5 10 15	
AAG GCT CGC TGC AGT CGG AAG GCA CTG CAT GTC AAC TTC AAG GAC ATG	937
Lys Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met	
20 25 30	
GGC TGG GAC GAC TGG ATC ATC GCA CCC CTT GAG TAC GAG GCT TTC CAC	985
Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His	
35 40 45	
TGC GAG GGG CTG TGC GAG TTC CCA TTG CGC TCC CAC CTG GAG CCC ACG	1033
Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr	
50 55 60	
AAT CAT GCA GTC ATC CAG ACC CTG ATG AAC TCC ATG GAC CCC GAG TCC	1081
Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser	
65 70 75	
ACA CCA CCC ACC TGC TGT GTG CCC ACG CGG CTG AGT CCC ATC AGC ATC	1129
Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile	
80 85 90 95	
CTC TTC ATT GAC TCT GCC AAC AAC GTG GTG TAT AAG CAG TAT GAG GAC	1177
Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp	
100 105 110	
ATG GTC GTG GAG TCG TGT GGC TGC AGG TAG	1207
Met Val Val Glu Ser Cys Gly Cys Arg	
115 120	

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ala Pro Leu Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys	
1 5 10 15	
Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly	
20 25 30	
Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys	
35 40 45	
Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn	
50 55 60	
His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr	
65 70 75 80	
Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu	
85 90 95	

Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met
 100 105 110
 Val Val Glu Ser Cys Gly Cys Arg
 115 120

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 128 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: V1-1 fragment
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 28..102

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGATCCTGGA AGGATTGGAT CATTGCG CCG CTG GAC TAC GAG GCG TAC CAC	51
Pro Leu Asp Tyr Glu Ala Tyr His	
1 5	
TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC	99
Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr	
10 15 20	
AAC CACGCTATAG TCCAAACCTT TCTAGA	128
Asn	
25	

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Leu Cys Asp Phe Pro	
1 5 10 15	
Leu Arg Ser His Leu Glu Pro Thr Asn	
20 25	

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 128 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo Sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: VL-1
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 28..102

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGATCCTGGG ATGACTGGAT TATGGCG CCG CTG GAC TAC GAG GCG TAC CAC	51
Pro Leu Asp Tyr Glu Ala Tyr His	
1 5	
TGC GAG GGT GTA TGC GAC TTC CCG CTG CGC TCG CAC CTG GAG CCC ACC	99
Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr	
10 15 20	
AAC CACGCCATGC TACAAACGCT TCTAGA	128
Asn	
25	

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro	
1 5 10 15	
Leu Arg Ser His Leu Glu Pro Thr Asn	
20 25	

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3585 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pALV1-781

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTAACTACCC AACTCAAAAA AAAAAAAAAA AAAAACCCCC TCTAACCCCC ATTGACGAAA	60
GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC	120
GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT	180

ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG	240
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCTT TTTTTCGGGC	300
ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA	360
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA	420
GAGTTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG	480
CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA TACACTATTC	540
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC	600
AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT	660
TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCCTTTTT TTGCACAACA TGGGGGATCA	720
TGTAAGTCGC CTTGATCGTT GGGAAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG	780
TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC AACTATTAA CTGGCGAACT	840
ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG	900
ACCACTTCTG CGCTCGGCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG	960
TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT	1020
CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC	1080
TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAAGTGTCA GACCAAGTTT ACTCATATAT	1140
ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT	1200
TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTTCG TTCCACTGAG CGTCAGACCC	1260
CGTAGAAAAG ATCAAAGGAT CTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT	1320
GCAAACAAAA AAACCACCGC TACCAGCGGT GGTGTGTTTG CCGGATCAAG AGCTACCAAC	1380
TCTTTTTCCG AAGGTAAGT GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT	1440
GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT	1500
GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA	1560
CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC	1620
ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG	1680
AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT	1740
CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC	1800
TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG	1860
GAGCCTATGG AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC	1920
TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC	1980
CTTTGAGTGA GCTGATACCG CTCGCCGAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG	2040
CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCCGCGCGTT GGCCGATTCA	2100
TTAATGCAGA ATTGATCTCT CACCTACCAA ACAATGCCCC CCTGCAAAAA ATAAATTCAT	2160
ATAAA--ACA TACAGATAAC CATCTGCGGT GATAAATTAT CTCTGGCGGT GTTGACATAA	2220

ATACCACTGG	CGGTGATACT	GAGCACATCA	GCAGGACGCA	CTGACCACCA	TGAAGGTGAC	2280
GCTCTTAAAA	ATTAAGCCCT	GAAGAAGGGC	AGCATTCAAA	GCAGAAGGCT	TTGGGGTGTG	2340
TGATACGAAA	CGAAGCATTG	GCCGTAAGTG	CGATTCCGGA	TTAGCTGCCA	ATGTGCCAAT	2400
CGCGGGGGGT	TTTCGTTTCTAG	GACTACAAC	GCCACACACC	ACCAAAGCTA	ACTGACAGGA	2460
GAATCCAGAT	GGATGCACAA	ACACGCCGCC	GCGAACGTCG	CGCAGAGAAA	CAGGCTCAAT	2520
GGAAAGCAGC	AAATCCCCTG	TTGGTTGGGG	TAAGCGCAAA	ACCAGTCCG	AAAGATTTTT	2580
TTAACTATAA	ACGCTGATGG	AAGCGTTTAT	GCGGAAGAGG	TAAAGCCCTT	CCCGAGTAAC	2640
AAAAAAACAA	CAGCATAAAT	AACCCCGCTC	TTACACATTC	CAGCCCTGAA	AAAGGGCATC	2700
AAATTAAACC	ACACCTATGG	TGTATGCATT	TATTTGCATA	CATTCAATCA	ATTGTTATCT	2760
AAGGAAATAC	TTACATATGT	CTCGTTGTTC	TCGTAAACCA	CTGCATGTAG	ATTTTAAAGA	2820
GCTCGGCTGG	GACGACTGGA	TCATCGCGCC	GCTGGACTAC	GAGGCGTACC	ACTGCGAGGG	2880
CCTTTGCGAC	TTCCCTTTGC	GTTGCGACCT	CGAGCCCACC	AACCATGCCA	TCATTGAGAC	2940
GCTGCTCAAC	TCCATGGCAC	CAGACGCGGC	GCCGGCCTCC	TGCTGTGTGC	CAGCGCGCCT	3000
CAGCCCCATC	AGCATCCTCT	ACATCGACGC	CGCCAACAAC	GTTGTCTACA	AGCAATACGA	3060
GGACATGGTG	GTGGAGGCCT	GCGGCTGCAG	GTAGTCTAGA	GTCGACCTGC	AGTAATCGTA	3120
CAGGGTAGTA	CAAATAAAAA	AGGCACGTCA	GATGACGTGC	CTTTTTTCTT	GTGAGCAGTA	3180
AGCTTGGCAC	TGGCCGTCGT	TTTACAACGT	CGTGACTGGG	AAAACCCTGG	CGTTACCCAA	3240
CTTAATCGCC	TTGCAGCACA	TCCCCCTTTC	GCCAGCTGGC	GTAATAGCGA	AGAGGCCCGC	3300
ACCGATCGCC	CTTCCCAACA	GTTGCGCAGC	CTGAATGGCG	AATGGCGCCT	GATGCGGTAT	3360
TTTCTCCTTA	CGCATCTGTG	CGGTATTTCA	CACCGCATAT	ATGGTGCACT	CTCAGTACAA	3420
TCTGCTCTGA	TGCCGCATAG	TTAAGCCAGC	CCCAGACACC	GCCAACACCC	GCTGACGCGC	3480
CCTGACGGGC	TTGTCTGCTC	CCGGCATCCG	CTTACAGACA	AGCTGTGACC	GTCTCCGGGA	3540
GCTGCATGTG	TCAGAGGTTT	TCACCGTCAT	CACCGAAACG	CGCGA		3585

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 272 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: mouse
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: mv1
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 28..243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCG CCA TTA GAC TAC	51
Trp Ile Ile Ala Pro Leu Asp Tyr	
1 5	
GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT CTG CGC TCG CAC	99
Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His	
10 15 20	
CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG	147
Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met	
25 30 35 40	
GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC GCA AGG CTC AGT	195
Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser	
45 50 55	
CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC GTG GTC TAC AAG	243
Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys	
60 65 70	
CAATACGAGG ACATGGTGGT GGGGAATTC	272

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 72 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val	
1 5 10 15	
Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile	
20 25 30	
Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser	
35 40 45	
Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp	
50 55 60	
Ala Ala Asn Asn Val Val Tyr Lys	
65 70	

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 272 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: mouse

(vii) IMMEDIATE SOURCE:

- (B) CLONE: mV2

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 28..243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATT ATC GCG CCC CTA GAG TAC      51
                    Trp Ile Ile Ala Pro Leu Glu Tyr
                      1             5

GAG GCC TAT CAC TGC GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC      99
Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His
   10             15             20

CTT GAG CCC ACT AAC CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG     147
Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Met Asn Ser Met
   25             30             35             40

GAC CCG GGC TCC ACC CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT     195
Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr
           45             50             55

CCC ATT AGC ATC CTG TAC ATC GAC GCG GGC AAT AAT GTA GTC TAC AAG     243
Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn Asn Val Val Tyr Lys
           60             65             70

CAATACGAGG ACATGGTGGT GGGGAATTC      272
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val
  1             5             10             15

Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile
   20             25             30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser
   35             40             45

Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp
   50             55             60

Ala Gly Asn Asn Val Val Tyr Lys
   65             70
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mouse

(vii) IMMEDIATE SOURCE:
(B) CLONE: mV9

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 28..243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCA CCT CTT GAG TAT      51
                  Trp Ile Ile Ala Pro Leu Glu Tyr
                   1                   5

GAG GCC TTC CAC TGC GAA GSA CTG TGT GAG TTC CCC TTG CGC TCC CAC      99
Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His
   10                   15                   20

TTG GAG CCC ACA AAC CAC GCA GTC ATT CAG ACC CTA ATG AAC TCT ATG     147
Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met
   25                   30                   35                   40

GAC CCT GAA TCC ACA CCA CCC ACT TGT TGT GTG CCT ACA CGG CTG AGT     195
Asp Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser
                   45                   50                   55

CCT ATT AGC ATC CTC TTC ATC GAC TCT GCC AAC AAC GTG GTG TAT AAA     243
Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys
   60                   65                   70

CAATACGAGG ACATGGTGGT GGGGAATTC      272
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu
  1                   5                   10                   15

Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val
   20                   25                   30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr
   35                   40                   45

Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp
   50                   55                   60

Ser Ala Asn Asn Val Val Tyr Lys
   65                   70
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: BMP/TGF-beta consensus sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Trp Xaa Asp Trp Ile Xaa Ala
1 5

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligonucleotide #1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGATCCTGG VANGAYTGGA THRINGC

27

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

(B) CLONE: BMP/TGF-beta consensus sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

His Ala Ile Xaa Gln Thr
1 5

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligonucleotide #2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
TTTCTAGAAR NGTYTGNACD ATNGCRTG

28

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: oligonucleotide #3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
CCACTGCGAG GGCCTTTGCG ACTTCCCTTT GCGTTCGCAC

40

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: oligonucleotide #4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
TGCGGATCCA GCCGCTGCAG CCGCAAGCC

29

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: oligonucleotide #5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
GACTCTAGAC TACCTGCAGC CGCAGGCCT

29

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide #6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCGGATCCAA GGAGCTCGGC TGGGACGA

28

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligonucleotide #7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGAATCCCC ACCACCATGT CCTCGTAT

28

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1171 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: Human VL-1 protein

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..964

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 605..964

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

G AAT TCG GAT CTC TCG CAC ACT CCT CTC CGG AGA CAG AAG TAT TTG
Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu
-201-200 -195 -190

46

TTT GAT GTG TCC ATG CTC TCA GAC AAA GAA GAG CTG GTG GGC GCG GAG
Phe Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu
-185 -180 -175

94

CTG CGG CTC TTT CGC CAG GCG CCC TCA GCG CCC TGG GGG CCA CCA GCC Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala -170 -165 -160 -155	142
GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG Gly Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu -150 -145 -140	190
GAC GCG CGG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu -135 -130 -125	238
GTC TTC GAC GTG TGG CAG GGC CTG CGC CAC CAG CCC TGG AAG CAG CTG Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu -120 -115 -110	286
TGC TTG GAG CTG CGG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG GCC Cys Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala -105 -100 -95	334
GAG GCG CGC GCG CGG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CGG Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg -90 -85 -80 -75	382
AGT CTG GGC TTC GGC CGG AGG GTG CGG CCT CCC CAG GAG CGG GCC CTG Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu -70 -65 -60	430
CTG GTG GTA TTC ACC AGA TCC CAG CGC AAG AAC CTG TTC GCA GAG ATG Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met -55 -50 -45	478
CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GGC GCC Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala -40 -35 -30	526
GAG GGG TCG TGG CCG CCG CCG TCG GGC GCC CCG GAT GCC AGG CCT TGG Glu Gly Ser Trp Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp -25 -20 -15	574
CTG CCC TCG CCC GGC CGC CGG CGG CGG CGC ACG GCC TTC GCC AGT CGC Leu Pro Ser Pro Gly Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg -10 -5 1 5	622
CAT GGC AAG CGG CAC GGC AAG AAG TCC AGG CTA CGC TGC AGC AAG AAG His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys 10 15 20	670
CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile 25 30 35	718
GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe 40 45 50	766
CCG CTG CGC TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACG Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr 55 60 65 70	814
CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val 75 80 85	862
CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GGC AAT Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn 90 95 100	910

AAT GTG GTC TAC AAG CAG TAC GAG GAC ATG GTG GTG GAG TCG TGC GGC	958
Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly	
105 110 115	
TGC AGG TAGCGGTGCC TTTCCCGCCG CCTTGGCCCG GAACCAAGGT GGGCCAAGGT	1014
Cys Arg	
120	
CCGCCTTGCA GGGGAGGCCT GGCTGCAGAG AGGCGGAGGA GGAAGCTGGC GCTGGGGGAG	1074
GCTGAGGGTG AGGGAACAGC CTGGATGTGA GAGCCGGTGG GAGAGAAGGG AGCGCACCTT	1134
CCCAGTAACT TCTACCTGCC AGCCCAGAGG GAAATAT	1171

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu Phe	
-201 -200 -195 -190	
Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu Leu	
-185 -180 -175 -170	
Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala Gly	
-165 -160 -155	
Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu Asp	
-150 -145 -140	
Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu Val	
-135 -130 -125	
Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu Cys	
-120 -115 -110	
Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala Glu	
-105 -100 -95 -90	
Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Asp Leu Arg Ser	
-85 -80 -75	
Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu	
-70 -65 -60	
Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met Arg	
-55 -50 -45	
Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala Glu	
-40 -35 -30	
Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp Leu	
-25 -20 -15 -10	
Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg His	
-5 1 5	
Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys Pro	
10 15 20	

Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala
 25 30 35
 Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro
 40 45 50 55
 Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu
 60 65 70
 Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val Pro
 75 80 85
 Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn Asn
 90 95 100
 Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys
 105 110 115
 Arg
 120

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: DNA encoding BMP2 propeptide/BMP-12 mature peptide

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1233

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 847..1233

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATG GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC	48
Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val	
-282 -280 -275 -270	
CTC CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG	96
Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys	
-265 -260 -255	
TTC GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG	144
Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu	
-250 -245 -240 -235	
GTC CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA	192
Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys	
-230 -225 -220	
CAG AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA	240
Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu	
-215 -210 -205	

GAC CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp -200 -195 -190	288
CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe -185 -180 -175	336
CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr -170 -165 -160 -155	384
ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe -150 -145 -140	432
ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala -135 -130 -125	480
TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile -120 -115 -110	528
ATA AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu -105 -100 -95	576
GAC ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp -90 -85 -80 -75	624
GTC ACC CCC GCT GTG ATG CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His -70 -65 -60	672
GGA TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser -55 -50 -45	720
AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser -40 -35 -30	768
TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys -25 -20 -15	816
GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT ACG GCG TTG GCC GGG ACG Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr -10 -5 1 5	864
CGG ACA GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC Arg Thr Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg 10 15 20	912
AGG GGC CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys 25 30 35	960
GAG CTC GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala 40 45 50	1008
TAC CAC TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu 55 60 65 70	1056

CCC ACC AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA	1104
Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro	
75 80 85	
GAC GCG GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC	1152
Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile	
90 95 100	
AGC ATC CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC	1200
Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr	
105 110 115	
GAG GAC ATG GTG GTG GAG GCC TGC GGC TGC AGG	1233
Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg	
120 125	

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val	
-282 -280 -275 -270	
Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys	
-265 -260 -255	
Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu	
-250 -245 -240 -235	
Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys	
-230 -225 -220	
Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu	
-215 -210 -205	
Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp	
-200 -195 -190	
His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe	
-185 -180 -175	
His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr	
-170 -165 -160 -155	
Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe	
-150 -145 -140	
Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala	
-135 -130 -125	
Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile	
-120 -115 -110	
Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu	
-105 -100 -95	
Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp	
-90 -85 -80 -75	

Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His
 -70 -65 -60
 Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser
 -55 -50 -45
 Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser
 -40 -35 -30
 Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys
 -25 -20 -15
 Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr
 -10 -5 1 5
 Arg Thr Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg
 10 15 20
 Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys
 25 30 35
 Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala
 40 45 50
 Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu
 55 60 65 70
 Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro
 75 80 85
 Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile
 90 95 100
 Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr
 105 110 115
 Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
 120 125

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1203 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: murine MV1
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 2..721

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

A AAG TTC TGC CTG GTG CTG GNG NCG GTG ACG GCC TCG GAG AGC AGN
 Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03
 1 5 10 15
 CNG CTG GCC CTG AGA CGA CTG GGC TTC GGC TGN CCG GGC GGT GGC GAC
 X04 Leu Ala Leu Arg Arg Leu Gly Phe Gly X05 Pro Gly Gly Gly Asp
 20 25 30

46

94

GGC GGC GGC ACT GCG GNC GAG GAG CGC GCG CTG TTG GTG ATC TCC TCC Gly Gly Gly Thr 35 Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser 40 45	142
CGT ACG CAA AGG AAA GAG AGT CTG TTC CGG GAG ATC CGA GCC CAG GCC Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala 50 55 60	190
CGT GCT CTC CGG GCC GCT GCA GAG CCG CCA CCG GAT CCA GGA CCA GGC Arg Ala Leu Arg Ala Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly 65 70 75	238
GCT GGG TCA CGC AAA GCC AAC CTG GGC GGT CGC AGG CGG CAG CGG ACT Ala Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr 80 85 90 95	286
GCG CTG GCT GGG ACT CGG GGA GNG NAG GGA AGC GGT GGT GGC GGC GGT Ala Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly 100 105 110	334
GGC GGT GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GCA Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala 115 120 125	382
GGC AGG GGC CAC GGG CGC AGA GGC CGG AGC CGC TGC GGT CGC AAG TCA Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser 130 135 140	430
CTG CAC GTG GAC TTT AAG GAG CTG GGC TGG GAC GAC TGG ATC ATC GCG Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp Trp Ile Ile Ala 145 150 155	478
CCA TTA GAC TAC GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro 160 165 170 175	526
CTG CGC TCG CAC CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu 180 185 190	574
CTC AAC TCC ATG GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro 195 200 205	622
GCA AGG CTC AGT CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn 210 215 220	670
GTG GTC TAC AAG CAG TAC GAA GAC ATG GTG GTG GAG GCC TGC GGC TGC Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys 225 230 235	718
AGG TAGCATGCGG TCTGGGGAGG GTCTGGCCGC CCAGGACCCT AGCTCAAGAG Arg 240	771
CAGGTGTCAT CAGGCCCCGAG GGACGGCGGA CTATGGCCTC TGCCAGCACA GAGGAGAGCA	831
CACAGTTAAC ACTCACATTT ACACACTCCT TCACTCACGC ACATGTTTAC CGTGGACGGC	891
AGGCGCTAAA AGCCTTGCTT ATTTGCTACC ATTGATACAA ACCTCTGTCC TTTTCGGGAG	951
AGGGAAGGGC ATCTGTGTTT ATGTTGCAGT AATTGGCACT AAATCCAAGT AGAAATGGGT	1011
TAGCATTGGA TTCTCCTTTT AGTTGGAGGC GGTGTGGCTG GATTCTGAC GTTGGATATG	1071
GAGTGCACTG CAGGGCTGGG ATACCCAGAT TCTCTGGAGT GGGCATTGGG AACCTTCAAA	1131

AGTAAGGAGC CACTGGGGCT TGGGAGGGAG CACCCGGTTC CTAAACAAGT CTGATGTGTA 1191
CTGCTCAGTT TG 1203

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 240 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Lys	Phe	Cys	Leu	Val	Leu	X01	X02	Val	Thr	Ala	Ser	Glu	Ser	X03	X04	
1				5					10					15		
Leu	Ala	Leu	Arg	Arg	Leu	Gly	Phe	Gly	X05	Pro	Gly	Gly	Gly	Asp	Gly	
			20					25						30		
Gly	Gly	Thr	Ala	X06	Glu	Glu	Arg	Ala	Leu	Leu	Val	Ile	Ser	Ser	Arg	
		35					40					45				
Thr	Gln	Arg	Lys	Glu	Ser	Leu	Phe	Arg	Glu	Ile	Arg	Ala	Gln	Ala	Arg	
	50					55					60					
Ala	Leu	Arg	Ala	Ala	Ala	Glu	Pro	Pro	Pro	Asp	Pro	Gly	Pro	Gly	Ala	
65				70						75					80	
Gly	Ser	Arg	Lys	Ala	Asn	Leu	Gly	Gly	Arg	Arg	Arg	Gln	Arg	Thr	Ala	
			85						90					95		
Leu	Ala	Gly	Thr	Arg	Gly	X07	X08	Gly	Ser	Gly	Gly	Gly	Gly	Gly	Gly	
		100				105								110		
Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Ala	Gly	
		115				120						125				
Arg	Gly	His	Gly	Arg	Arg	Gly	Arg	Ser	Arg	Cys	Gly	Arg	Lys	Ser	Leu	
	130					135					140					
His	Val	Asp	Phe	Lys	Glu	Leu	Gly	Trp	Asp	Asp	Trp	Ile	Ile	Ala	Pro	
145					150					155					160	
Leu	Asp	Tyr	Glu	Ala	Tyr	His	Cys	Glu	Gly	Val	Cys	Asp	Phe	Pro	Leu	
			165					170						175		
Arg	Ser	His	Leu	Glu	Pro	Thr	Asn	His	Ala	Ile	Ile	Gln	Thr	Leu	Leu	
			180					185					190			
Asn	Ser	Met	Ala	Pro	Asp	Ala	Ala	Pro	Ala	Ser	Cys	Cys	Val	Pro	Ala	
		195				200						205				
Arg	Leu	Ser	Pro	Ile	Ser	Ile	Leu	Tyr	Ile	Asp	Ala	Ala	Asn	Asn	Val	
	210					215				220						
Val	Tyr	Lys	Gln	Tyr	Glu	Asp	Met	Val	Val	Glu	Ala	Cys	Gly	Cys	Arg	
225					230					235					240	

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1046 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: MURINE MV2

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..790

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

A AGA AAA CAA GCT TGC ATT CCT GCA GGT CCG ACT CTA AGA GGA TCC	46
Arg Lys Gln Ala Cys Ile Pro Ala Gly Pro Thr Leu Arg Gly Ser	
1 5 10 15	
TCA GGG ACC CAA CCC AGG CCG GCT GGG AAG TCT TTC GAC GTG TGG CAG	94
Ser Gly Thr Gln Pro Arg Pro Ala Gly Lys Ser Phe Asp Val Trp Gln	
20 25 30	
GGC CTG CGC CCT CAG CCT TGG AAG CAG CTG TGC CTG GAG TTG CGG GCA	142
Gly Leu Arg Pro Gln Pro Trp Lys Gln Leu Cys Leu Glu Leu Arg Ala	
35 40 45	
GCC TGG GGT GAG CTG GAC RCC GGG GAT ACG GGG GCG CGC GCG AGG GGT	190
Ala Trp Gly Glu Leu Asp X01 Gly Asp Thr Gly Ala Arg Ala Arg Gly	
50 55 60	
CCC CAG CAG CCA CCG CCT CTG GAC CTG CGG AGT CTG GGC TTC GGT CGG	238
Pro Gln Gln Pro Pro Pro Leu Asp Leu Arg Ser Leu Gly Phe Gly Arg	
65 70 75	
AGG GTG AGA CCG CCC CAG GAG CGC GCC CTG CTT GTA GTG TTC ACC AGA	286
Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg	
80 85 90 95	
TCG CAG CGC AAG AAC CTG TTC ACT GAG ATG CAT GAG CAG CTG GGC TCT	334
Ser Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser	
100 105 110	
GCA GAG GCT GCG GGA GCC GAG GGG TCA TGT CCA GCG CCG TCG GGC TCC	382
Ala Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser	
115 120 125	
CCA GAC ACC GGG TCT TGG CTG CCC TCG CCC GGC CGC CGG CGG CGA CGC	430
Pro Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg	
130 135 140	
ACC GCC TTC GCC AGC CGT CAC GGC AAG CGA CAT GGC AAG AAG TCC AGG	478
Thr Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg	
145 150 155	
CTG CGC TGC AGC AGA AAG CCT CTG CAC GTG AAT TTT AAG GAG TTA GGC	526
Leu Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly	
160 165 170 175	
TGG GAC GAC TGG ATT ATC GCG CCC CTA GAG TAC GAG GCC TAT CAC TGC	574
Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys	
180 185 190	

GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC CTT GAG CCC ACT AAC	622
Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn	
195 200 205	
CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG GAC CCG GGC TCC ACC	670
His Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr	
210 215 220	
CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT CCC ATT AGC ATC CTG	718
Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu	
225 230 235	
TAC ATC GAC GCG GGC AAT AAT GTN GTC TAC AAG CAG TAT GAG GAC ATG	766
Tyr Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met	
240 245 250 255	
GTG GTG GAG TCC TGC GGC TGT AGG TAGCGGTGCT GTCCCGCCAC CTGGGCCAGG	820
Val Val Glu Ser Cys Gly Cys Arg	
260	
GACCATGGAG GGAGGCCTGA CTGCCGAGAA AGGAGCAGGA GCTGGCCTTG GAAGAGGCCA	880
CAGGTGGGGG ACAGCCTGAA AGTAGGAGCA CAGTAAGAAG CAGCCCAGCC TTCCCAGAAC	940
CTTCCAATCC CCCAACCAG AAGCAGCTAA GGGGTTTCAC AACTTTTGGC CTTGCCAGCC	1000
TGGAAGACT AGACAAGAGG GATTCTTCTC TTTTATTAT GGCTTG	1046

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Arg Lys Gln Ala Cys Ile Pro Ala Gly Pro Thr Leu Arg Gly Ser Ser	
1 5 10 15	
Gly Thr Gln Pro Arg Pro Ala Gly Lys Ser Phe Asp Val Trp Gln Gly	
20 25 30	
Leu Arg Pro Gln Pro Trp Lys Gln Leu Cys Leu Glu Leu Arg Ala Ala	
35 40 45	
Trp Gly Glu Leu Asp X01 Gly Asp Thr Gly Ala Arg Ala Arg Gly Pro	
50 55 60	
Gln Gln Pro Pro Pro Leu Asp Leu Arg Ser Leu Gly Phe Gly Arg Arg	
65 70 75 80	
Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg Ser	
85 90 95	
Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser Ala	
100 105 110	
Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser Pro	
115 120 125	
Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg Thr	
130 135 140	

Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg Leu
145 150 155 160

Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp
165 170 175

Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu
180 185 190

Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His
195 200 205

Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro
210 215 220

Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr
225 230 235 240

Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met Val
245 250 255

Val Glu Ser Cys Gly Cys Arg
260

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1345 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:
(B) CLONE: HUMAN V1-1

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 138..1301

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 990..1301

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AACTATAGCA CCTGCAGTCC CTGGTCTTGG GTGTAGGGGT GCGCTCCTGG TCCCGCGGCT	60
CAGGGATATG CAGTGACCAA TGGGTTGTTG GCCTGATGGG ACTTTTGGCT TGCTAAACCA	120
AACTCGGTT CGGATAG CCC GGG CGA AGA CGT CCG CTG CTC TGG GCC AGG	170
Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg	
-284 -280 -275	
CTG GCA GCG TTC AGG CTG GGG CAG AGA CGC GGA GTC GGG CGC TGG CTC	218
Leu Ala Ala Phe Arg Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu	
-270 -265 -260	

CAA CAG GCC TGG CTC CCA CAT CGA AGA CAG CTG GGC CAT TTG CTG TTA Gln Gln Ala Trp Leu Pro His Arg Arg Gln Leu Gly His Leu Leu Leu -255 -250 -245	266
GGA GGC CCC GCG CTG ACA GTG TGC AGG ATT TGC TCT TAC ACA GCT CTT Gly Gly Pro Ala Leu Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu -240 -235 -230	314
TCT CTC TGT CCC TGC CGG TCC CCC GCA GAC GAA TCG GCA GCC GAA ACA Ser Leu Cys Pro Cys Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr -225 -220 -215 -210	362
GGC CAG AGC TTC CTG TTC GAC GTG TCC AGC CTT AAC GAC GCA GAC GAG Gly Gln Ser Phe Leu Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu -205 -200 -195	410
GTG GTG GGT GCC GAG CTG CGC GTG CTG CGC CGG GGA TCT CCA GAG TCG Val Val Gly Ala Glu Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser -190 -185 -180	458
GGC CCA GGC AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC Gly Pro Gly Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys -175 -170 -165	506
CCG GGC GCC GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG Pro Gly Ala Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu -160 -155 -150	554
CCC CTA GTC GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG Pro Leu Val Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met -145 -140 -135 -130	602
AGG CGC CAC CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG Arg Arg His Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu -125 -120 -115	650
CTG CGC GCA GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA Leu Arg Ala Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg -110 -105 -100	698
CTG GGC TTC GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC Leu Gly Phe Gly Trp Pro Gly Gly Gly Ser Ala Ala Glu Glu Arg -95 -90 -85	746
GCG GTG CTA GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC Ala Val Leu Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe -80 -75 -70	794
CGG GAG ATC CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA Arg Glu Ile Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser -65 -60 -55 -50	842
GAG CCG CTG CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC Glu Pro Leu Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val -45 -40 -35	890
ATT GGC GGC CGC AGA CCG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA Ile Gly Gly Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr -30 -25 -20	938
GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly -15 -10 -5	986
CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu 5 10 15	1034

GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC	1082
Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His	
20 25 30	
TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC	1130
Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr	
35 40 45	
AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG	1178
Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala	
50 55 60	
GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC	1226
Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile	
65 70 75	
CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC	1274
Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp	
80 85 90 95	
ATG GTG GTG GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG	1321
Met Val Val Glu Ala Cys Gly Cys Arg	
100	
GGGCAGCCAC GCGGCCGAGG ATCC	1345

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 388 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg Leu Ala Ala Phe Arg	
-284 -280 -275 -270	
Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu Gln Gln Ala Trp Leu	
-265 -260 -255	
Pro His Arg Arg Gln Leu Gly His Leu Leu Leu Gly Gly Pro Ala Leu	
-250 -245 -240	
Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu Ser Leu Cys Pro Cys	
-235 -230 -225	
Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr Gly Gln Ser Phe Leu	
-220 -215 -210 -205	
Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu Val Val Gly Ala Glu	
-200 -195 -190	
Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser Gly Pro Gly Ser Trp	
-185 -180 -175	
Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala Ala Arg	
-170 -165 -160	
Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val Gly Gln	
-155 -150 -145	
Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His Arg Arg	
-140 -135 -130 -125	

